IN THE CLAIMS:

Please amend Claims 1 to 64, 67 to 69, 71 to 85, 87 to 94 and 97 to 100, and add new Claims 101 and 102, as shown below. The claims, as pending in the subject application, read as follows:

1. (Currently Amended) A ~~feature~~ comparison apparatus comprising:

~~input means for receiving~~ a receiver operable to receive an input signal;

a recognition ~~processing means for comparing~~ processor operable to compare said input signal with stored ~~feature~~ label models to ~~identify an input~~ generate a recognised sequence of ~~features~~ labels in said input signal and confidence data representative of the confidence that the ~~input~~ recognised sequence of ~~features~~ labels is representative of said input signal; and

~~means for comparing~~ a similarity measure calculator operable to compare said ~~input~~ recognised sequence of ~~features~~ labels with a stored sequence of ~~features~~ labels using a linear weighted combination of predetermined ~~similarity~~ confusion data which defines ~~similarities~~ confusability between different ~~features~~ labels and ~~using~~ said confidence data, to provide a measure of the similarity between the ~~input~~ recognised sequence of ~~features~~ labels and the stored sequence of ~~features~~ labels.

2. (Currently Amended) An apparatus according to claim 1, wherein confidence data is stored for said stored sequence of ~~features~~ labels and wherein said ~~comparing means provides~~ calculator is operable to calculate said similarity measure using, in addition, said stored confidence data.

3. (Currently Amended) ~~The~~ An apparatus according to claim 1, wherein said confidence data comprises confidence data associated with each ~~feature~~ label in the recognised sequence of ~~features~~ labels.

4. (Currently Amended) An apparatus according to claim 1, wherein said recognition ~~processing means~~ processor is operable to output a list of alternatives for each ~~feature~~ label in said ~~input~~ recognised sequence of ~~features~~ labels and confidence data associated with each alternative.

5. (Currently Amended) An apparatus according to claim 1, wherein said ~~comparing means~~ similarity measure calculator comprises:

~~means for aligning features~~ an aligner operable to align labels of the ~~input~~ recognised sequence of ~~features~~ labels with ~~features~~ labels of the stored sequence of ~~features~~ labels to form a number of aligned pairs of ~~features~~ labels;

~~means for comparing~~ a comparator operable to compare the ~~features~~ labels of each aligned pair of ~~features~~ labels using said weighted combination of said predetermined ~~similarity~~ confusion data and said confidence data, to generate a comparison score representative of the similarity between the aligned pair of ~~features~~ labels; and

~~means for combining~~ a combiner operable to combine the comparison scores for all the aligned pairs of ~~features~~ labels to provide said similarity measure.

Sub B1

6. (Currently Amended) An apparatus according to claim 5, wherein said ~~comparing means~~ comparator comprises:

a first ~~comparing means for comparing~~ sub-comparator operable to compare, for each aligned pair, the ~~input~~ recognised sequence ~~feature~~ label in the aligned pair with each of a plurality of ~~features~~ labels taken from a set of predetermined ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a corresponding plurality of intermediate comparison scores representative of the similarity between said ~~first~~ recognised sequence ~~feature~~ label and the respective ~~features~~ labels from the set;

A6

a second ~~comparing means for comparing~~ sub-comparator operable to compare, for each aligned pair, the stored sequence ~~feature~~ label in the aligned pair with each of said plurality of ~~features~~ labels from the set using said ~~similarity~~ confusion data and said confidence data to provide a further corresponding plurality of intermediate comparison scores representative of the similarity between said stored sequence ~~feature~~ label and the respective ~~features~~ labels from the set; and

~~means for calculating~~ a calculator operable to calculate said comparison score for the aligned pair by combining said pluralities of intermediate comparison scores.

7. (Currently Amended) An apparatus according to claim 6, wherein said first and second ~~comparing means~~ sub-comparators are operable to compare the ~~input~~ recognised sequence ~~feature~~ label and the stored sequence ~~feature~~ label respectively with each of the ~~features~~ labels in said set of predetermined ~~features~~ labels.

8. (Currently Amended) An apparatus according to claim 6, wherein said ~~comparing means~~ comparator is operable to generate a comparison score for an aligned pair of ~~features~~ labels which represents a probability of confusing the stored sequence ~~feature~~ label of the aligned pair as the ~~input~~ recognised sequence ~~feature~~ label of the aligned pair.

9. (Currently Amended) An apparatus according to claim 8, wherein said first and second ~~comparing means~~ sub-comparators are operable to provide intermediate comparison scores which are indicative of a probability of confusing the corresponding ~~feature~~ label taken from the set of predetermined ~~features~~ labels as the ~~feature~~ label in the aligned pair.

10. (Currently Amended) An apparatus according to claim 9, wherein said ~~calculating means~~ calculator is operable (i) to multiply the intermediate comparison scores obtained when comparing the ~~input~~ recognised and stored sequence ~~features~~ labels in the aligned pair with the same ~~feature~~ label from the set to provide a plurality of multiplied intermediate comparison scores; and (ii) to add the resulting multiplied intermediate comparison scores, to calculate said comparison score for the aligned pair.

11. (Currently Amended) An apparatus according to claim 10, wherein each of said ~~features~~ labels in said set of predetermined ~~features~~ labels has a predetermined probability of occurring within a sequence of ~~features~~ labels and wherein said ~~calculating~~

~~means~~ calculator is operable to ~~weigh~~ weight each of said multiplied intermediate comparison scores with the respective probability of occurrence for the ~~feature~~ label from the set used to generate the multiplied intermediate comparison scores.

12. (Currently Amended) An apparatus according to claim 11, wherein said ~~calculating means~~ calculator is operable to calculate:

$$\sum_{r=1}^{n} P(q_j|p_r)P(a_i|p_r)P(p_r)$$

where q_j and a_i are an aligned pair of ~~input~~ recognised and stored sequence ~~features~~ labels respectively; $P(q_j|p_r)$ is the probability of confusing set ~~feature~~ label p_r as ~~input~~ recognised sequence ~~feature~~ label q_j; $P(a_i|p_r)$ is the probability of confusing set ~~feature~~ label p_r as stored sequence ~~feature~~ label a_i; and $P(p_r)$ represents the probability of set ~~feature~~ label p_r occurring in a sequence of ~~features~~ labels.

13. (Currently Amended) An apparatus according to claim 12, wherein the confusion probabilities for the ~~input~~ recognised and stored sequence ~~features~~ labels are determined in advance and depend upon the recognition system that was used to generate the respective ~~input in~~ recognised and stored sequences.

14. (Currently Amended) An apparatus according to claim 12, wherein said ~~calculating means~~ calculator is operable to calculate $P(q_j|p_r)$ using said ~~similarity~~ confusion data and said confidence data.

15. (Currently Amended) An apparatus according to claim 14, wherein said ~~calculating means~~ calculator is operable to take a weighted combination of said ~~similarity~~ confusion data and said confidence data to determined $P(q_j|p_r)$.

16. (Currently Amended) An apparatus according to claim 15, wherein said ~~similarity~~ confusion data is obtained from a training session in which a large amount of input signals for which the ~~feature~~ label content is known are processed by said recognition ~~processing means~~ processor and wherein said ~~calculating means~~ calculator is operable to ~~weigh~~ weight said confidence data for a current ~~feature~~ label in dependence upon the amount of training data available for the current ~~feature~~ label in said training session.

17. (Currently Amended) An apparatus according to claim 16, wherein said ~~calculating means~~ calculator is operable to calculate:

$$P(q_j|p_r) = \frac{b_c c_{jr} + b_e e^r_{qj} + \beta}{b_c n_r + b_e + N_p \beta}$$

where c_{jr} and n_r are counts generated during the training session for the number of times the recognition ~~processing means~~ processor decoded ~~feature~~ label q_j when it should have decoded ~~feature~~ label p_r and the number of times the recognition ~~processing means~~ processor decoded anything when it should have decoded ~~feature~~ label p_r, respectively; $e_{qj}{}^r$ is the confidence data associated with the ~~input~~ recognised sequence ~~feature~~ label of the aligned pair which is associated with the set ~~feature~~ label P_r; β

represents a lower limit of the confidence probabilities; N_p is the total number of ~~features~~ labels in the set; and b_c and b_e are scaling factors which are applied to the ~~similarity~~ confusion data and the confidence data respectively.

18. (Currently Amended) An apparatus according to claim ~~claims~~ 10, wherein said intermediate comparison scores represent log probabilities and wherein said ~~calculating means~~ calculator is operable to perform said multiplication by adding the respective intermediate comparison scores and is operable to perform said addition of said multiplied scores by performing a log addition calculation.

19. (Currently Amended) An apparatus according to claim 18, wherein said ~~combining means~~ combiner is operable to add the comparison scores for all the aligned pairs to determine said similarity measure.

20. (Currently Amended) An apparatus according to claim 5, wherein said ~~aligning means~~ aligner is operable to identify ~~feature~~ label deletions and insertions in said ~~input~~ recognised and stored sequences of ~~features~~ labels and wherein said ~~comparing means~~ comparator is operable to generate said comparison score for an aligned pair of ~~features~~ labels in dependence upon ~~feature~~ label deletions and insertions identified by said ~~aligning means~~ aligner which occur in the vicinity of the ~~features~~ labels in the aligned pair.

21. (Currently Amended) An apparatus according to claim 5, wherein said ~~aligning means comprises dynamic programming means for aligning~~ aligner is operable to align said ~~input~~ recognised and stored sequences of ~~features~~ labels using a dynamic programming technique.

22. (Currently Amended) An apparatus according to claim 21, wherein said ~~dynamic programming means~~ aligner is operable to determine progressively a plurality of possible alignments between said ~~input~~ recognised and stored sequences of ~~features~~ labels and wherein said ~~comparing means~~ comparator is operable to determine a comparison score for each of the possible aligned pairs of ~~features~~ labels determined by said ~~dynamic programming means~~ aligner.

23. (Currently Amended) An apparatus according to claim 22, wherein said ~~comparing means~~ comparator is operable to generate said comparison score during the progressive determination of said possible alignments.

24. (Currently Amended) An apparatus according to claim 21, wherein said ~~dynamic programming means~~ aligner is operable to determine an optimum alignment between said ~~input~~ recognised and said stored sequences of ~~features~~ labels and wherein said ~~combining means~~ combiner is operable to provide said similarity measure by combining the comparison scores only for the optimum aligned pairs of ~~features~~ labels.

25. (Currently Amended) An apparatus according to claim 22, wherein said ~~combining means~~ combiner is operable to provide said similarity measure by combining all the comparison scores for all the possible aligned pairs of ~~features~~ labels.

26. (Currently Amended) An apparatus according to ~~any of claims~~ claim 6, wherein each of the ~~features~~ labels in said ~~input~~ recognised and stored sequences of ~~features~~ labels belong to said set of predetermined ~~features~~ labels.

27. (Currently Amended) An apparatus according to claim 26, wherein said ~~similarity~~ confusion data comprises, for each ~~feature~~ label in the set of ~~features~~ labels, the number of times the recognition ~~processing means~~ processor decodes the ~~feature~~ label when it should have decoded a different ~~feature~~ label, for each different ~~feature~~ label, and the number of times the recognition ~~processing means~~ processor decodes anything when it should have decoded the different ~~feature~~ label, for each different ~~feature~~ label.

28. (Currently Amended) An apparatus according to claim 27, wherein said predetermined confusion data further comprises, for each ~~feature~~ label in the set of ~~features~~ labels, the probability of inserting the ~~feature~~ label in a sequence of ~~features~~ labels.

29. (Currently Amended) An apparatus according to claim 26, wherein said predetermined confusion data further comprises, for each ~~feature~~ label in the set of

~~features~~ labels, the probability of deleting the ~~feature~~ label from a sequence of ~~features~~ labels.

30. (Currently Amended) An apparatus according to claim 1, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent time sequential signals.

31. (Currently Amended) An apparatus according to claim 1, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent audio signals.

32. (Currently Amended) An apparatus according to claim 31, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent speech.

33. (Currently Amended) An apparatus according to claim 32, wherein each of said ~~features~~ labels represents a sub-word unit of speech.

34. (Currently Amended) An apparatus according to claim 33, wherein each of said ~~features~~ labels represents a phoneme.

35. (Currently Amended) An apparatus according to claim 1, wherein said ~~input~~ recognised sequence of ~~features~~ labels comprises a plurality of sub-word units

generated from a typed input and wherein said ~~similarity~~ confusion data comprises mis-typing probabilities and/or mis-spelling probabilities.

36. (Currently Amended) An apparatus according to claim 1, wherein said stored sequence of ~~features~~ labels comprises a sequence of sub-word units generated from a spoken input and wherein said ~~similarity~~ confusion data comprises mis-recognition probabilities.

37. (Currently Amended) An apparatus according to claim 1, wherein said ~~comparing means~~ comparator is operable to compare said ~~input~~ recognised sequence of ~~features~~ labels with a plurality of stored sequences of ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a respective measure of the similarity between the ~~input~~ recognised sequence of ~~features~~ labels and said plurality of stored sequences of ~~features~~ labels.

38. (Currently Amended) An apparatus according to claim 37, further comprising ~~means for comparing~~ a similarity measure comparator operable to compare said plurality of similarity measures output by said ~~comparing means~~ comparator and ~~means for outputting~~ operable to output a signal indicative of the stored sequence of ~~features~~ labels which is most similar to said ~~input~~ recognised sequence of ~~features~~ labels.

39. (Currently Amended) An apparatus according to claim 37, wherein said ~~comparing means~~ comparator comprises ~~normalising means for normalising~~ a normaliser operable to normalise each of said similarity measures.

40. (Currently Amended) An apparatus according to claim 39, wherein said ~~normalising means~~ normaliser is operable to normalise each similarity measure by dividing each similarity measure by a respective normalisation score which varies in dependence upon the length of the corresponding stored sequence of ~~features~~ labels.

41. (Currently Amended) An apparatus according to claim 40, wherein the respective normalisation scores vary in dependence upon the sequence of ~~features~~ labels in the corresponding stored sequence of ~~features~~ labels.

AC

42. (Currently Amended) An apparatus according to claim 40, wherein said ~~comparing means~~ comparator comprises:

a first ~~comparing means for comparing~~ sub-comparator operable to compare, for each aligned pair, the ~~input~~ recognised sequence ~~feature~~ label in the aligned pair with each of a plurality of ~~features~~ labels taken from a set of predetermined ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a corresponding plurality of intermediate comparison scores representative of the similarity between said ~~first~~ recognised sequence ~~feature~~ label and the respective ~~features~~ labels from the set;

a second ~~comparing means for comparing~~ sub-comparator operable to compare, for each aligned pair, the stored sequence ~~feature~~ label in the aligned pair with each of said plurality of ~~features~~ labels from the set using said ~~similarity~~ confusion data and said confidence data to provide a further corresponding plurality of intermediate comparison scores representative of the similarity between said stored sequence ~~feature~~ label and the respective ~~features~~ labels from the set; and

~~means for calculating~~ a calculator operable to calculate said comparison score for the aligned pair by combining said pluralities of intermediate comparison scores; and

wherein said respective normalisation scores vary with the corresponding intermediate comparison scores calculated by said second ~~comparing means~~ sub-comparator.

Ao

43. (Currently Amended) An apparatus according to claim 42, wherein said ~~aligning means comprises dynamic programming means for aligning~~ aligner is operable to align said ~~input~~ recognised and stored sequences of ~~features~~ labels using a dynamic programming technique and wherein said ~~normalising means~~ normaliser is operable to calculate the respective normalisation scores during the progressive calculation of said possible alignments by said ~~dynamic programming means~~ aligner.

Sub B1

44. (Currently Amended) An apparatus according to claim 43, wherein said ~~normalising means~~ normaliser is operable to calculate, for each possible aligned pair of ~~features~~ labels:

$$\sum_{r=1}^{n} P(a_i|p_r)P(p_r)$$

where $P(a_i|p_r)$ represents the probability of confusing set ~~feature~~ label p_r as stored sequence ~~feature~~ label a_i and $P(p_r)$ represents the probability of set ~~feature~~ label p_r occurring in a sequence of ~~features~~ labels.

45. (Currently Amended) An apparatus according to claim 44, wherein said ~~normalising means~~ normaliser is operable to calculate said respective normalisations by multiplying the normalisation terms calculated for the respective aligned pairs of ~~features~~ labels.

Ae

46. (Currently Amended) An apparatus for searching a database comprising a plurality of information entries to identify information to be retrieved therefrom, each of said plurality of information entries having an associated annotation comprising a sequence of ~~features~~ labels, the apparatus comprising:

a ~~feature~~ label comparison apparatus according to claim 1 for comparing a query sequence of ~~features~~ labels obtained from an input query with the ~~features~~ labels of each annotation to provide a set of comparison results; and

~~means for identifying~~ an information identifier operable to identify said information to be retrieved from said database using said comparison results.

47. (Currently Amended) An apparatus according to claim 46, wherein said ~~feature~~ label comparison apparatus has a plurality of different comparison modes of operation and ~~in that~~ wherein the apparatus further comprises:

~~means for determining~~ a determiner operable to determine if the sequence of ~~features~~ labels of a current annotation was generated from an audio signal or from text, and for outputting a determination result; and

~~means for selecting~~ a selector operable to select, for the current annotation, the mode of operation of said ~~feature~~ label comparison apparatus in dependence upon said determination result.

Ae

48. (Currently Amended) A ~~feature~~ label comparison method comprising ~~the steps of~~:

receiving an input signal;

a recognition processing step of comparing said input signal with stored ~~feature~~ label models to ~~identify an input~~ generate a recognised sequence of ~~features~~ labels in said input signal and confidence data representative of the confidence that the ~~input~~ recognised sequence of ~~features~~ labels is representative of said input signal; and

calculating a measure of the similarity between the recognised sequence of labels and a stored sequence of labels by comparing said ~~input~~ recognised sequence of

~~features~~ labels with [a] the stored sequence of ~~features~~ labels using a linear weighted combination of predetermined ~~similarity~~ confusion data which defines ~~similarities~~ confusability between different ~~features~~ labels and ~~using~~ said confidence data~~, to provide a measure of the similarity between the input sequence of features and the stored sequence of features~~.

49. (Currently Amended) A method according to claim 48, wherein confidence data is stored for said stored sequence of ~~features~~ labels and wherein said ~~comparing~~ calculating step ~~provides~~ calculates said similarity measure using, in addition, said stored confidence data.

50. (Currently Amended) A method according to claim 48, wherein said confidence data comprises confidence data associated with each ~~feature~~ label in the sequence of ~~features~~ labels.

51. (Currently Amended) A method according to claim 48, wherein said recognition processing step outputs a list of alternatives for each ~~feature~~ label in said ~~input~~ recognised sequence of ~~features~~ labels and confidence data associated with each alternative.

52. (Currently Amended) A method according to claim 48, wherein said ~~comparing~~ calculating step comprises the steps of:

aligning ~~features~~ labels of the ~~input~~ recognised sequence of ~~features~~ labels with ~~features~~ labels of the stored sequence of ~~features~~ labels to form a number of aligned pairs of ~~features~~ labels;

comparing the ~~features~~ labels of each aligned pair of ~~features~~ labels using said predetermined ~~similarity~~ confusion data and said confidence data, to generate a comparison score representative of the similarity between the aligned pair of ~~features~~ labels; and

combining the comparison scores for all the aligned pairs of ~~features~~ labels to provide said similarity measure.

53. (Currently Amended) A method according to claim 52, wherein said comparing step comprises:

a first ~~comparing~~ sub-comparing step of comparing, for each aligned pair, the ~~input~~ recognised sequence ~~feature~~ label in the aligned pair with each of a plurality of ~~features~~ labels taken from a set of predetermined ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a corresponding plurality of intermediate comparison scores representative of the similarity between said ~~first~~ recognised sequence ~~feature~~ label and the respective ~~features~~ labels from the set;

a second ~~comparing~~ sub-comparing step of comparing, for each aligned pair, the stored sequence ~~feature~~ label in the aligned pair with each of said plurality of ~~features~~ labels from the set using said ~~similarity~~ confusion data and said confidence data to provide a further corresponding plurality of intermediate comparison scores representative

Ae

of the similarity between said stored sequence ~~feature~~ label and the respective ~~features~~ labels from the set; and

a step of calculating said comparison score for the aligned pair by combining said pluralities of intermediate comparison scores.

54. (Currently Amended) A method according to claim 53, wherein said first and second ~~comparing~~ sub-comparing steps compare the ~~input~~ recognised sequence ~~feature~~ label and the stored sequence ~~feature~~ label respectively with each of the ~~features~~ labels in said set of predetermined ~~features~~ labels.

55. (Currently Amended) A method according to claim 53, wherein said comparing step generates a comparison score for an aligned pair of ~~features~~ labels which represents a probability of confusing the stored sequence ~~feature~~ label of the aligned pair as the ~~input~~ recognised sequence ~~feature~~ label of the aligned pair.

AC

56. (Currently Amended) A method according to claim 55, wherein said first and second ~~comparing~~ sub-comparing steps provide intermediate comparison scores which are indicative of a probability of confusing the corresponding ~~feature~~ label taken from the set of predetermined ~~features~~ labels as the ~~feature~~ label in the aligned pair.

57. (Currently Amended) A method according to claim 56, wherein said calculating step (i) multiplies the intermediate scores obtained when comparing the ~~input~~

recognised and stored sequence ~~features~~ labels in the aligned pair with the same ~~feature~~ label from the set to provide a plurality of multiplied intermediate comparison scores; and (ii) adds the resulting multiplied intermediate scores, to calculate said comparison score for the aligned pair.

58. (Currently Amended) A method according to claim 57, wherein each of said ~~features~~ labels in said set of predetermined ~~features~~ labels has a predetermined probability of occurring within a sequence of ~~features~~ labels and wherein said calculating step ~~weighs~~ weights each of said multiplied intermediate comparison scores with the respective probability of occurrence for the ~~feature~~ label from the set used to generate the multiplied intermediate comparison scores.

59. (Currently Amended) A method according to claim 58, wherein said calculating step calculates:

KL

$$\sum_{r=1}^{n} P(q_j|p_r)P(a_i|p_r)P(p_r)$$

where q_j and a_i are an aligned pair of ~~input~~ recognised and stored sequence ~~features~~ labels respectively; $P(q_j|p_r)$ is the probability of confusing set ~~feature~~ label p_r as ~~input~~ recognised sequence ~~feature~~ label q_j; $P(a_i|p_r)$ is the probability of confusing set ~~feature~~ label p_r as stored sequence ~~feature~~ label a_i; and $P(p_r)$ represents the probability of set ~~feature~~ label p_r occurring in a sequence of ~~features~~ labels.

60. (Currently Amended) A method according to claim 59, wherein the confusion probabilities for the ~~input~~ recognised and stored sequence ~~features~~ labels are determined in advance and depend upon the recognition system that was used to generate the respective ~~input~~ recognised and stored sequences.

61. (Currently Amended) A method according to claim 59, wherein said calculating step calculates $P(q_j|p_r)$ using said ~~similarity~~ confusion data and said confidence data.

62. (Currently Amended) A method according to claim 61, wherein said calculating step takes a weighted combination of said ~~similarity~~ confusion data and said confidence data to ~~determined~~ determine $P(q_j|p_r)$.

Ae

63. (Currently Amended) A method according to claim 62, wherein said ~~similarity~~ confusion data is obtained from a training session in which a large amount of input signals for which the ~~feature~~ label content is known are processed by said recognition processing step and wherein said calculating step ~~weighs~~ weights said confidence data for a current ~~feature~~ label in dependence upon the amount of training data available for the current ~~feature~~ label in said training session.

64. (Currently Amended) A method according to claim 63, wherein said calculating step calculates:

$$P(q_j|p_r) = \frac{b_c c_{jr} + b_e e_{qj}^r + \beta}{b_c n_r + b_e + N_p \beta}$$

where c_{jr} and n_r are counts generated during the training session for the number of times the recognition processing step decoded ~~feature~~ label q_j when it should have decoded ~~feature~~ label p_r and the number of times the recognition processing step decoded anything when it should have decoded ~~feature~~ label p_r, respectively; e_{qj}^r is the confidence data associated with the ~~input~~ recognised sequence ~~feature~~ label of the aligned pair which is associated with the set ~~feature~~ label P_r; β represents a lower limit of the confidence probabilities; N_p is the total number of ~~features~~ labels in the set; and b_c and b_e are scaling factors which are applied to the ~~similarity~~ confusion data and the confidence data respectively.

65. (Original) A method according to claim 57, wherein said intermediate scores represent log probabilities and wherein said calculating step performs said multiplication by adding the respective intermediate scores and performs said addition of said multiplied scores by performing a log addition calculation.

66. (Original) A method according to claim 65, wherein said combining step adds the comparison scores for all the aligned pairs to determine said ~~similarity~~ confusion measure.

67. (Currently Amended) A method according to claim 52, wherein said aligning step identifies ~~feature~~ label deletions and insertions in said ~~input~~ recognised and stored sequences of ~~features~~ labels and wherein said comparing step generates said comparison score for an aligned pair of ~~features~~ labels in dependence upon ~~feature~~ label deletions and insertions identified by said aligning step which occur in the vicinity of the ~~features~~ labels in the aligned pair.

68. (Currently Amended) A method according to claim 52, wherein said aligning step comprises a dynamic programming step for aligning said ~~input~~ recognised and stored sequences of ~~features~~ labels using a dynamic programming technique.

69. (Currently Amended) A method according to claim 68, wherein said dynamic programming step progressively determines a plurality of possible alignments between said ~~input~~ recognised and stored sequences of ~~features~~ labels and wherein said comparing step determines a comparison score for each of the possible aligned pairs of ~~features~~ labels determined by said dynamic programming step.

70. (Original) A method according to claim 69, wherein said comparing step generates said comparison score during the progressive determination of said possible alignments.

71. (Currently Amended) A method according to claim 68, wherein said dynamic programming step determines an optimum alignment between said ~~input~~ recognised and said stored sequences of ~~features~~ labels and wherein said combining step provides said similarity measure by combining the comparison scores only for the optimum aligned pairs of ~~features~~ labels.

72. (Currently Amended) A method according to claim 69, wherein said combining step provides said similarity measure by combining all the comparison scores for all the possible aligned pairs of ~~features~~ labels.

73. (Currently Amended) A method according to claim 53, wherein each of the ~~features~~ labels in said ~~input~~ recognised and stored sequences of ~~features~~ labels belong to said set of predetermined ~~features~~ labels.

Al

74. (Currently Amended) A method according to claim 73, wherein said ~~similarity~~ confusion data comprises, for each ~~feature~~ label in the set of ~~features~~ labels, the number of times the recognition processing step decodes the ~~feature~~ label when it should have decoded a different ~~feature~~ label, for each different ~~feature~~ label, and the number of times the recognition processing step decodes anything when it should have decoded the different ~~feature~~ label, for each different ~~feature~~ label.

75. (Currently Amended) A method according to claim 74, wherein said predetermined data further comprises, for each ~~feature~~ label in the set of ~~features~~ labels, the probability of inserting the ~~feature~~ label in a sequence of ~~features~~ labels.

76. (Currently Amended) A method according to claim 74, wherein said predetermined data further comprises, for each ~~feature~~ label in the set of ~~features~~ labels, the probability of deleting the ~~feature~~ label from a sequence of ~~features~~ labels.

77. (Currently Amended) A method according to claim 48, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent time sequential signals.

78. (Currently Amended) A method according to claim 48, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent audio signals.

79. (Currently Amended) A method according to claim 78, wherein said ~~input~~ recognised and stored sequences of ~~features~~ labels represent speech.

80. (Currently Amended) A method according to claim 79, wherein each of said ~~features~~ labels represents a sub-word unit of speech.

81. (Currently Amended) A method according to claim 80, wherein each of said ~~features~~ labels represents a phoneme.

82. (Currently Amended) A method according to claim 48, wherein said ~~input~~ recognised sequence of ~~features~~ labels comprises a plurality of sub-word units generated from a typed input and wherein said ~~similarity~~ confusion data comprises mis-typing probabilities and/or mis-spelling probabilities.

83. (Currently Amended) A method according to claim 48, wherein said stored sequence of ~~features~~ labels comprises a sequence of sub-word units generated from a spoken input and wherein said ~~similarity information~~ confusion data comprises mis-recognition probabilities.

84. (Currently Amended) A method according to claim 48, wherein said comparing step compares said ~~input~~ recognised sequence of ~~features~~ labels with a plurality of stored sequences of ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a respective measure of the similarity between the ~~input~~ recognised sequence of ~~features~~ labels and said plurality of stored sequences of ~~features~~ labels.

85. (Currently Amended) A method according to claim 84, further comprising the step of comparing said plurality of similarity measures output by said

comparing step and the step of outputting a signal indicative of the stored sequence of ~~features~~ labels which is most similar to said ~~input~~ recognised sequence of ~~features~~ labels.

86. (Original) A method according to claim 84, wherein said comparing step comprises a normalising step for normalising each of said ~~similarity~~ confusion measures.

87. (Currently Amended) A method according to claim 86, wherein said normalising step normalises each similarity measure by dividing each similarity measure by a respective normalisation score which varies in dependence upon the length of the corresponding stored sequence of ~~features~~ labels.

88. (Currently Amended) A method according to claim 87, wherein the respective normalisation scores vary in dependence upon the sequence of ~~features~~ labels in the corresponding stored sequence of ~~features~~ labels.

89. (Currently Amended) A method according to claim 87, wherein said comparing step comprises:

a first ~~comparing~~ sub-comparing step of comparing, for each aligned pair, the ~~input~~ recognised sequence ~~feature~~ label in the aligned pair with each of a plurality of ~~features~~ labels taken from a set of predetermined ~~features~~ labels using said ~~similarity~~ confusion data and said confidence data to provide a corresponding plurality of

intermediate comparison scores representative of the similarity between said ~~first~~ recognised sequence ~~feature~~ label and the respective ~~features~~ labels from the set;

a second ~~comparing~~ sub-comparing step of comparing, for each aligned pair, the stored sequence ~~feature~~ label in the aligned pair with each of said plurality of ~~features~~ labels from the set using said ~~similarity~~ confusion data and said confidence data to provide a further corresponding plurality of intermediate comparison scores representative of the similarity between said stored sequence ~~feature~~ label and the respective ~~features~~ labels from the set; and

a step of calculating said comparison score for the aligned pair by combining said pluralities of intermediate comparison scores; and

wherein said respective normalisation scores vary with the corresponding intermediate comparison scores calculated by said second ~~comparing~~ sub-comparing step.

90. (Currently Amended) A method according to claim 86, wherein said aligning step comprises a dynamic programming step for aligning said ~~input~~ recognised and stored sequences of ~~features~~ labels using a dynamic programming technique and wherein said normalising step calculates the respective normalisation scores during the progressive calculation of said possible alignments by said dynamic programming step.

91. (Currently Amended) A method according to claim 90, wherein said normalising step calculates, for each possible aligned pair of ~~features~~ labels:

$$\sum_{r=1}^{n} P(a_i|p_r)P(p_r)$$

where $P(a_i|p_r)$ represents the probability of confusing set ~~feature~~ label p_r as stored sequence ~~feature~~ label a_i and $P(p_r)$ represents the probability of set ~~feature~~ label p_r occurring in a sequence of ~~features~~ labels.

92. (Currently Amended) A method according to claim 91, wherein said normalising step calculates said respective normalisations by multiplying the normalisation terms calculated for the respective aligned pairs of ~~features~~ labels.

93. (Currently Amended) A method of searching a database comprising a plurality of information entries to identify information to be retrieved therefrom, each of said plurality of information entries having an associated annotation comprising a sequence of ~~features~~ labels, the method comprising:

a ~~feature~~ label comparison method according to claim 48 for comparing a query sequence of ~~features~~ labels obtained from an input query with the ~~features~~ labels of each annotation to provide a set of comparison results; and

a step of identifying said information to be retrieved from said database using said comparison results.

94. (Currently Amended) A method according to claim 93, wherein said ~~feature~~ label comparison method has a plurality of different comparison

modes of operation and in that the method further comprises the steps of:

determining if the sequence of ~~features~~ labels of a current annotation was generated from an audio signal or from text, and outputting a determination result; and

selecting, for the current annotation, the mode of operation of said ~~feature~~ label comparison method in dependence upon said determination result.

95. (Original) A method according to claim 93, wherein one or more of said information entries is the associated annotation.

96. (Original) A method according to claim 48, wherein the method steps are performed in the order in which they are claimed.

97. (Currently Amended) A storage medium storing processor implementable instructions for controlling a processor to implement a ~~feature~~ comparison method, the process steps comprising steps for:

receiving an input signal;

comparing the input signal with stored ~~feature~~ label models to ~~identify an input~~ generate a recognised sequence of ~~features~~ labels in said input signal and confidence data representative of the confidence that the ~~input~~ recognised sequence of ~~features~~ labels is representative of the input signal; and

calculating a measure of the similarity between the recognised sequence of labels and a stored sequence of labels by comparing the ~~input~~ recognised sequence of

~~features~~ labels with [a] the stored sequence of ~~features~~ labels using a linear weighted combination of predetermined ~~similarity~~ confusion data which defines ~~similarities~~ confusability between different ~~features~~ labels and ~~using~~ said confidence data~~, to provide a measure of the similarity confusion between the input sequence of features and the stored sequence of features~~.

98. (Currently Amended) A storage medium storing processor implementable instructions for controlling a processor to implement a method of searching a database comprising a plurality of information entries to identify information to be retrieved therefrom, each of the plurality of information entries having an associated annotation comprising a sequence of ~~features~~ labels, the process steps comprising:

the process steps stored on the medium according to claim 97 for comparing a query sequence of ~~features~~ labels obtained from an input query with the ~~features~~ labels of each annotation to provide a set of comparison results; and

~~a step of identifying~~ process steps to identify said information to be retrieved from said database using said comparison results.

99. (Currently Amended) Processor implementable instructions for controlling a processor to implement a ~~feature~~ comparison method, the process steps comprising steps for:

receiving an input signal;

comparing the input signal with stored ~~feature~~ label models to ~~identify an input~~ generate a recognised sequence of ~~features~~ labels in said input signal and confidence data representative of the confidence that the ~~input~~ recognised sequence of ~~features~~ labels is representative of the input signal; and

calculating a measure of the similarity between the recognised sequence of labels and a stored sequence of labels by comparing the ~~input~~ recognised sequence of ~~features~~ labels with [a] the stored sequence of ~~features~~ labels using a linear weighted combination of predetermined ~~similarity~~ confusion data which defines ~~similarities~~ confusability between different ~~features~~ labels and ~~using~~ said confidence data~~, to provide a measure of the similarity between the input sequence of features and the stored sequence of features~~.

100. (Currently Amended) Processor implementable instructions for controlling a processor to implement a method of searching a database comprising a plurality of information entries to identify information to be retrieved therefrom, each of the plurality of information entries having an ~~a~~ associated annotation comprising a sequence of ~~features~~ labels, the process steps comprising:

the process steps of claim 99 for comparing a query sequence of ~~features~~ labels obtained from an input query with the ~~features~~ labels of each annotation to provide a set of comparison results; and

~~a step of identifying~~ process steps to identify said information to be retrieved from said database using said comparison results.

101. (New) A comparison apparatus comprising:

means for receiving an input signal;

recognition processing means for comparing said input signal with stored label models to generate a recognised sequence of labels in said input signal and confidence data representative of the confidence that the recognised sequence of labels is representation of the input signal; and

means for comparing said recognised sequence of labels with a stored sequence of labels using a linear weighted combination of predetermined confusion data which defines confusability between difference labels and said confidence data, to provide a measure of the similarity between the recognised sequence of features and the stored sequence of features.

102. (New) An apparatus for searching a database comprising a plurality of information entries to identify information to be retrieved therefrom, each of said plurality of information entries having an associated annotation comprising a sequence of labels, the apparatus comprising:

a comparison apparatus according to claim 101 for comparing a query sequence of labels obtained from an input query with the labels of each annotation to provide a set of comparison results; and

means for identifying said information to be retrieved from said database using said comparison results.